

JP-A-61-107397

In voice recognition response device, the device prompts a user for a proper voice input if the user's voice input is improper. For example, the device produces a message "please speak slowly" if the user speaks too fast for voice recognition.

6. 利用者の発声のしかたが不適切である場合  
不適切の内容を検出して、訂正用の音声応答する

## 資料④

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A)

昭61-107397

⑬ Int. Cl.<sup>4</sup>

識別記号

庁内整理番号

⑭ 公開 昭和61年(1986)5月26日

G 10 L 3/00

7350-5D

審査請求 未請求 発明の数 1 (全5頁)

⑮ 発明の名称 音声認識応答装置

⑯ 特 願 昭59-229248

⑰ 出 願 昭59(1984)10月31日

⑱ 発 明 者 上 原 堅 助 青梅市末広町2丁目9番地 株式会社東芝青梅工場内  
⑲ 出 願 人 株 式 会 社 東 芝 川崎市幸区堀川町72番地  
⑳ 代 理 人 弁 理 士 鈴 江 武 彦 外2名

### 明 細 書

#### 1. 発明の名称

音声認識応答装置

#### 2. 特許請求の範囲

電話回線を通じて送信される音声信号に対して音声認識処理に必要な音声区間を検出しかつその音声区間の検出が不可能な場合には検出不可内容を含む検出不可信号を出力する音声検出手段と、この音声検出手段で検出された音声区間の上記音声信号に対する音声認識処理を行なう音声認識手段と、この音声認識手段の認識結果に対応する音声応答処理または上記音声検出手段から出力される検出不可信号に対して予め決定された音声応答処理を実行して各音声応答処理による音声応答信号をそれぞれ上記電話回線に出力する音声応答手段とを具備してなることを特徴とする音声認識応答装置。

#### 3. 発明の詳細な説明

[発明の技術分野]

本発明は、電話回線を利用して音声認識応答処

理を実行する音声認識応答装置に関する。

[発明の技術的背景とその問題点]

音声認識応答システムは、電話回線を利用して、利用者とホストコンピュータ間で情報を交換し、利用者に対して音声による情報サービスを提供するシステムである。このシステムでは、第2図に示すように、利用者は電話器10を操作して、例えば予め決定された一連のコード番号を一路毎に区切って連続して発声する。利用者の発声による音声信号は、電話回線11を通じて音声認識装置13に送信される。このとき、回線切換回路12はホストコンピュータ14の制御により、電話回線11を音声認識装置13側に接続させる。音声認識装置13は、不特定話者単語認識用の音声認識装置からなり、利用者から発声された各数字を単語として一路毎に音声認識処理する。

ここで、利用者は、通常では音声応答装置15から送信される助誘音(ビー音)Pに応じて、必要な音声(音声信号S)を発声するようになっている。このとき、ホストコンピュータ14の制御によ

特開昭61-107397(2)

り回線切換回路12が音声応答装置15側に接続されて、音声応答装置15から出力される勧誘音Pが電話回線11を通じて利用者に送信される。勧誘音Pの出力が終了後に、第3図に示すような利用者からの音声信号Sが音声認識装置13に出力される。音声認識装置13は、音声信号Sから予め決定された閾値しを越える音声区間T1を検出し、この音声区間T1に対する認識処理を行なう。

ところで、利用者の中には上記のような勧誘音Pを聞き終わらないうちに、音声を発声する場合がある。この場合、第4図に示すように、音声信号Sの始端部分が音声認識装置13に入力されない事象が発生する。このため、音声認識装置13では、利用者の音声を誤認識またはリジェクトする場合が多くなる。このような利用者は、特に数字を連続して発声する場合に、たびたび早めに発声する。さらに、発声する単語または利用者の発声の癖により、発声が通常より極端に長くなる場合がある。音声認識装置13は認識処理を行なう際に一時バッファメモリに音声信号Sを格納するが、発声期間

の長い音声信号Sはそのバッファメモリの容量を越えることがある。このため、音声認識装置13では、第5図に示すように、本来の音声区間T1の始端部分が削除された音声区間T2に対して認識処理が実行されることになる。このため、利用者の発声に対して、誤認識またはリジェクトが発生することになる。また、利用者に中には音声を入力する際に、長時間に渡り沈黙している場合がある。従来のシステムでは、音声応答装置15からの勧誘音Pが出力した後、適当に時間(タイムアウト時間T3と称し、通常5秒程度)内に音声入力がない場合、利用者に対して発声を催促するように構成されている。即ち、例えば音声応答装置15から、「音声の入力がありませんので、もう一度どうぞ」等のメッセージを利用者に送信する。ここで、利用者がタイムアウト時間T3の限度付近で発声した場合、第6図に示すように、音声信号Sの始端部分が削除されて音声認識装置13に入力される場合がある。この場合でも、音声信号に対する誤認識またはリジェクトが発生することにな

る。

上記のように、従来のシステムでは、利用者の発声が不適切である場合には、音声認識率が著しく低下し、結果的に確実な音声応答サービスが不可能になる問題がある。

#### 〔発明の目的〕

本発明の目的は、音声認識応答システムにおいて、利用者の発声が不適切である場合、利用者に対して適切な発声を示唆して、利用者からの音声に対する確実な音声認識処理を実現し、音声応答サービスを確実に行なうことができる音声認識応答装置を提供することにある。

#### 〔発明の概要〕

本発明では、音声認識応答システムにおいて、利用者の不適切な発声を検出する音声検出手段が設けられている。この音声検出手段は、音声信号に対して音声認識処理に必要な音声区間を検出し、かつその音声区間の検出が不可能な場合には検出不可内容を含む検出不可信号を出力する。上記音声検出手段から検出不可信号が出力されると、音

声応答手段は予め決定された音声応答処理を実行し、音声応答処理による音声応答信号を電話回線を通じて利用者に送信するように構成されている。

このような構成により、利用者は発声が不適切な場合には、音声応答手段からの指示に応じて再度適切な発声を行なうことが可能となる。これにより、利用者が発声した音声に対する音声認識処理を確実に行なうことができる。

#### 〔発明の実施例〕

以下図面を参照して本発明の一実施例を説明する。第1図は同実施例に係わる音声認識応答装置の構成を示すブロック図である。第1図において、音声検出回路16は、回線切換回路12により電話回線11に接続されると、この電話回線11を通じて入力される音声信号Sに対する音声区間の検出処理を行なう。音声検出回路16は、利用者の不適切な発声により音声区間の検出が不可能な場合には、検出不可内容を含む検出不可信号Rをホストコンピュータ14に出力する。ホストコンピュータ14は、音声検出回路16からの検出不可信号Rに応じて、

音声応答装置15に予め決定された応答データを出力する。音声応答装置15は、応答データに応じた音声応答信号を電話回線11に出力する。

一方、音声検出回路16で検出された音声区間に対応する音声信号Sは、音声認識装置13に出力されて音声認識処理が実行される。ホストコンピュータ14は、音声認識装置13の認識結果に応じた応答データを音声応答装置15に出力する。

このような構成の音声認識応答装置において、図1の実例の動作を説明する。まず、音声認識応答システムに電話回線11が接続されて、システムが動作状態になると、ホストコンピュータ14は動詞音Pの発生を行なうために、音声応答装置15の動作を開始する。このとき、回線切換回路12はホストコンピュータ14の制御により、音声応答装置15側に電話回線11を接続させている。音声応答装置15から出力された動詞音Pは、電話回線11を通じて利用者に送信される。ホストコンピュータ14は、動詞音Pの発生終了後直ちに、電話回線11が音声検出回路16側に接続されるように回線切換回路12

の動作を開始する。

利用者は、動詞音Pに応じて、必要な音声を発声して、音声信号Sを電話回線11を通じて音声検出回路16に inputs する。音声検出回路16は、通常では音声信号Sに対する音声区間を検出して、その区間の音声信号Sを一時バッファメモリに格納した後に音声認識装置13に出力する。ここで、音声区間とは予め決定された閾値しより高レベルになった時を始端とし、また閾値しより低レベルになった時を終端とする区間である（第3図を参照）。音声認識装置13は、音声検出回路16で検出された音声区間Iに対応する音声信号Sの認識処理を実行し、認識結果をホストコンピュータ14に出力する。ホストコンピュータ14は、音声認識装置13からの認識結果に応じた応答データを音声応答装置15に出力する。即ち、例えば利用者からのコード番号に対して、ホストコンピュータ14は預金残高等の応答データを出力する。音声応答装置15は、応答データに応じた音声応答信号を利用者に送信することになる。

ところで、利用者が、上記のような動詞音Pを送信中に音声が発声したとする。この場合、回線切換回路12が電話回線11を音声検出回路16側に接続する前に、音声信号Sが入力されるため、その始端部分が削除された状態の音声信号Sが音声検出回路16に inputs される（第4図を参照）。音声検出回路16は、音声信号Sの始端が第4図に示すような閾値しより大きいレベルLaであることを検出すると、上記のような不適切な発声であると判定する。これにより、音声検出回路16は、音声信号Sの終端を検出した後に検出不可信号Rをホストコンピュータ14に出力する。ホストコンピュータ14は、音声検出回路16からの検出不可信号Rに基づいて、利用者の不適切な発声を訂正させるために予め決定された応答データを音声応答装置15に出力する。音声応答装置15は、応答データに応じて、例えば「音声の発声が早すぎます。信号音の後に発声して下さい」等のメッセージからなる音声応答信号を電話回線11に出力する。さらに、利用者の発声期間が長く、音声信号Sが音声検出

回路16のバッファメモリの容量を超える場合（第5図を参照）には、音声検出回路16は音声信号Sの終端を検出できないため検出不可信号Rをホストコンピュータ14に出力する。ホストコンピュータ14は、そのときの検出不可信号Rの内容に基づいて、利用者の不適切な発声を訂正させるために予め決定された応答データを音声応答装置15に出力する。音声応答装置15は、例えば「もう少し発声を短くして下さい」等のメッセージの音声応答信号を電話回線11に出力する。また、動詞音Pの発生後、利用者の沈黙が長く、音声検出回路16においてタイムアウト時間T3を過ぎても音声信号Sの終端が検出されない場合（第6図を参照）、音声検出回路16は検出不可信号Rをホストコンピュータ14に出力する。これにより、上記と同様にホストコンピュータ14からの応答データに応じて、音声応答装置15は例えば「音声の発声が遅すぎます。信号音の後にすぐには発声して下さい」等のメッセージの音声応答信号を電話回線11に出力する。このようにして、利用者の不適切な発声により、

(4)

利用者からの音声信号Sから確実に音声区間を検出できない場合、不適切な発声の内容に応じた音声応答信号が利用者に送信される。このため、利用者は不適切な発声を訂正して、再度適切な発声で音声信号を入力することができる。したがって、音声認識装置13には適正な音声区間の音声信号が入力されることになり、利用者からの音声信号に対する認識率を大幅に高めることができる。

〔発明の効果〕

以上詳述したように本発明によれば、音声認識応答システムにおいて、利用者の発声が不適切である場合、不適切の内容を検出して、その内容に応じた発声訂正用の音声応答信号を出力することができる。したがって、利用者に対して適切な発声で再度音声信号を入力することを催促することができ、音声信号に対する認識処理を確実に行うことが可能となる。これにより、結果的に利用者に対する音声応答サービスを、常に確実に提供することができるものである。

4、図面の簡単な説明

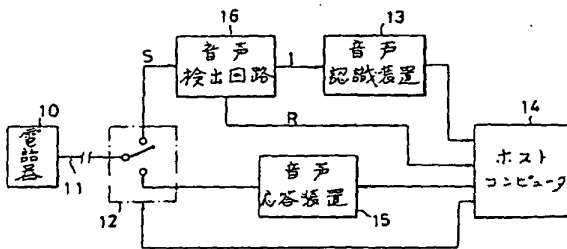
特開昭61-107397(4)

第1図は本発明の一実施例に係わる音声認識応答装置の構成を示すブロック図、第2図は従来の音声認識応答装置の構成を示すブロック図、第3図乃至第6図はそれぞれ本発明の動作を説明するための音声信号の波形図である。

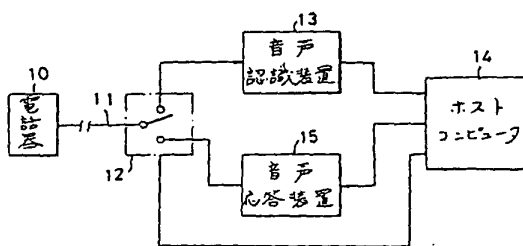
11…電話回線、12…回路切換回路、13…音声認識装置、14…ホストコンピュータ、15…音声応答装置、16…音声検出回路。

出願人代理人 弁理士 船江武彦

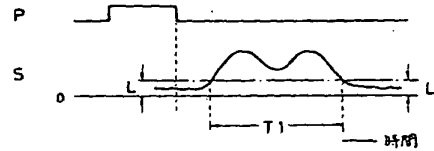
第1図



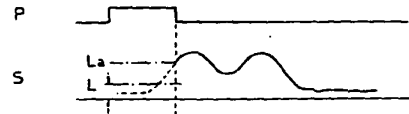
第2図



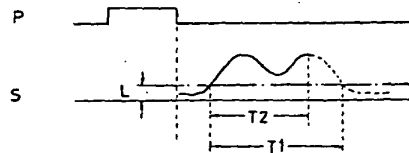
第3図



第4図



第5図



(5)

特開昭61-107397(5)

第 6 図

